

papla, cDNA and translated amino acid sequence:

I C F L L D D S T T F R N G T L N R G M 60
ATTGCTTCTTGTAGATGATTCGACGACATTTAGAAATGGTACCTTGAATAGAGGCATG
S P D C T F N E K D I V F Y V Y S R D K 120
TCTCCGGATTGTACTTTTAATGAGAAAGATATAGTATTCTATGTTTACTCAAGGATAAG
R D G I I L K K E T L T N Y D L F T K S 180
CGAGATGGTATTATTCTTAAGAAAGAAACTTTAACGAATTACGATCTGTTACAAAAGTCT
T I S K Q V V F L I H G F L S T G N N E 240
ACAATATCAAAACAAGTTGTATTCTTATACATGGTTTCCTTTCACACTGGGAATAATGAA
N F V A M S K A L I E K D D F L V I S V 300
AACTTCGTTGCTATGTCGAAAGCTTTAATAGAAAAAGATGATTTTCTTGTAATTTCGGTC
D W K K G A C N A F A S T K D A L G Y S 360
GACTGGAAGAGGGTGCTTGTAATGCTTTTGCTTCAACAAAGGATGCTTTGGGTTATTCC
K A V G N T R H V G K F V A D F T K L L 420
AAAGCCGTTGGAAACACACACGTCACGTTGGAAATTTGTAGCTGATTTTACAAAACACTACTT
V E K Y K V L I S N I R L I G H S L G A 480
GTAGAAAATATAAAGTGCTGATATCAAAATATACGATTGATCGGCATAGTTTGGGCGCG
H T S G F A G K E V Q K L K L G K Y K E 540
CATACTTCAGGTTTTGCGGGAAAGAAGTTCAAAAGTTAAAAATTAGGAAATACAAAGGAA

A A

Figure 1A

A . - - - - - A

I I G L D P A G P Y F H R S D C P D R L
ATTATCGGGCTTGATCCTGCTGGACCGTATTTTCATCGGAGTGACTGTCCGGACAGACTT 600

C V T D A E Y V Q V I H T S I I L G V Y
TGCCTAACAGACGCAGAAATATGTTCAAGTTATACATACATCAATCATATTAGGAGTATAT 660

Y N V G S V D F Y V N Y G K N Q P G C N
TATAATGTTGGTAGCGTTGATTCTACGTGAATTATGGAAAAAATCAACCTGGTTGCAAT 720

E P S C S H T K A V K Y L T E C I K H E
GAACCATCCTGCTCTCATACGAAAGCCGTGAAATATCTGACTGAGTGCATAAAACATGAA 780

C C L I G T P W K K Y F S T P K P I S Q
TGTTGTTTAAATTGGAACACCATGGAAGAAATATTTTCAGCACTCCAAAACCAATTTCCAG 840

C R G D T C V C V G L N A K S Y P A R G
TGCAGAGGAGACACCTGTGTTTGC GTTGATTGAATGC AAAAGTTATCCTGTAGAGGC 900

A F Y A P V E A N A P Y C H N E G I K L
GCATTTTATGCACCGGTTGAAGCAAATGCACCTTATTGCCATAACGAGGGGATTAAACTT 960

*
TAATTATAACAAAAGTCAATGTACACAAAATGTATCTATTGATGAATATTAAATGAAT 1020

AAACGAACAGTCAAATAAAAAA AAAA 1048

Note: The amino acid sequence ICFL.....GTLNR represents a portion of the leader sequence, as
Venom protein has the sequence of GMSPD.....

Figure 1B

papla, intron 1, (between nucleotides 111-112; see papla file):

AGGTAATAATCTCGATTCTATGCGTACGCGATTTTGTGATTATTTTCAAGAAAATGTA	60
AGAAAAATTTTAAAAATATATTACTGAAGTATGAAATAAAAACTTTTATACTTT	114

Figure 2A

papla, intron 2, (between nucleotides 720-721; see papla file):

GGTAATATTTTATATATAAATGAACAATTCTATGGAATAGAAATAGTACAAGCATCGAT	60
TATATCCTATGCCCTGTATATGATTTCGGAGTTAGACACTATTATTTTAAATAATTTT	120
TACATTA	127

Figure 2B

Vespid plas:

- wfh, white face hornet (D. maculate); vv, yellow jacket (V. vulgaris); pa, wasp (P. annularis):

1						50
wfh	~~FSVCPFSN	DTVKMIFLTR	ENRK.HDFYT	LDTMNRHNEF		KKSIKRPVV
vv	~~GPKCPEFS	DTVSHIETR	ENRN.RDLYT	LQTLQNHPEF		KKKTITRPVV
pa	GMSPDCTFNE	KDIVFYVYSR	DKRDGIILKK	E.TLTNYDLF		TKSTISKQVV
51						100
wfh	F.ITHGFTSS	ATEKNFVAMS	EALMHTGDFL	IIMVDWRMAA		CTDEYPGLKY
vv	F.ITHGFTSS	ASETNFINLA	KALVDKDNYM	VISIDWQTAA		CTNEAAGLKY
pa	FLI.HGFLST	GNNEFVAMS	KALIEKDDFL	VISVDWKKGA		C.NAFASTKD
101						150
wfh	.MFYKAAVGN	TRLVGNFIAM	IAKKLVEQYK	VPMTNIRLVG		HSLGAHISGF
vv	.LYYPTAARN	TRLVGQYIAT	ITQKLVKHYK	ISMANIRLLIG		HSLGAHISGF
pa	ALGYSKAVGN	TRHVGKFVAD	FTKLLVEKYK	VLISNIRLLIG		HSLGAHISGF
151						200
wfh	AGKRVQELKL	GKFSEIIGLD	PAGPSFKKND	CSERICETDA		HYVQILHTSS
vv	AGKRVQELKL	GKFSEIIGLD	PARPSFDSNH	CSERLCETDA		EYVQIIHTSN
pa	AGKRVQELKL	GKFKEIIGLD	PAGPYFHRSD	CPDRLCVTDA		EYVQVIHTSI

A-----A

Figure 3A

A - - - - - A

	201				251
wfh	NLGTERTLGT	VDFYINNGSN	QPGCRYIIGE	TCSHTRAVKY	FTECIRRECC
vv	TLGTEKTLGT	VDFYMNNGKN	QPGCGRFFSE	VCSHSRAVIY	MAECIKHECC
pa	ILGVYYNVGS	VDFYVNYGKN	QPGCNEPS..	.CSHTKAVKY	LTECIKHECC
	251				300
wfh	LIGVPQSK..	.SPQPVSKCT	RNECVCVGLN	AKKYPKRGSF	YVPVEAEAPY
vv	LIGPKSK..	.SSQPISSCT	KQECVCVGLN	AKKYPSRGSF	YVPVESTAPF
pa	LIGTPWKKYF	STPKPISQCR	GDTVCVCVGLN	AKSYPARGAF	YAPVEANAPY
	301				
wfh	CNNNGKII				
vv	CNNKGKII				
pa	CHNEGIKL				

Figure 3B

Pahya, cDNA and translated amino acid sequence:

Y V S L S P D S V F N	
TATGTGTCATTGTCCCCCGACTCAGTATTTAA	480
I I T D D I S H Q I L S R S N C E R S K	
TATCATCACCGATGACATCTCCACCAAAATTCCTTCCAGATCGAATTGTGAAAGATCCAA	540
R P K R V F S I Y W N V P T F N C H Q Y	
AAGACCGAAAGGGTCTTCAGCATTTTATTGGAAACGTTCTACCTTTATGTGCCACCAATA	600
G M N F D E V T D F N I K H N S K D N F	
TGGCATGAATTTTCGACGAGGTGACAGATTTTAATATCAAAACATAATTCTAAGGACAAATTT	660
R G E T I S I Y Y D P G K F P A L M P L	
TCGCGGTGAAACTATATCAATTTATTACGATCCTGGAAAAATTTCCAGCATTGATGCCACT	720
K N G N Y E E R N G G V P Q R G N I T I	
AAAAAATGGTAATTATGAGGAAAGAAACGGAGGGTTCTCTCAGCGAGGTAACATCACGAT	780
H L Q Q F N E D L D K M T P D K N F G G	
ACATTTGCAACAATTTAACGAAGATTTGGATAAAATGACACCGGATAAAAAATTTCCGGTGG	840
I G V I D F E R W K P I F R Q N W G N T	
TATCGGTGTAATTCGAAAGATGGAAACCGATTTTCCGACAGAAATTTGGGGTAACAC	900
E I H K K Y S I E L V R K E H P K W S E	
GGAATACATAAGAAATATTCTATTGAACTCGTTTCGGAAAGAACATCCAAAGTGGAGCGA	960

A . - - - - - A

Figure 4A

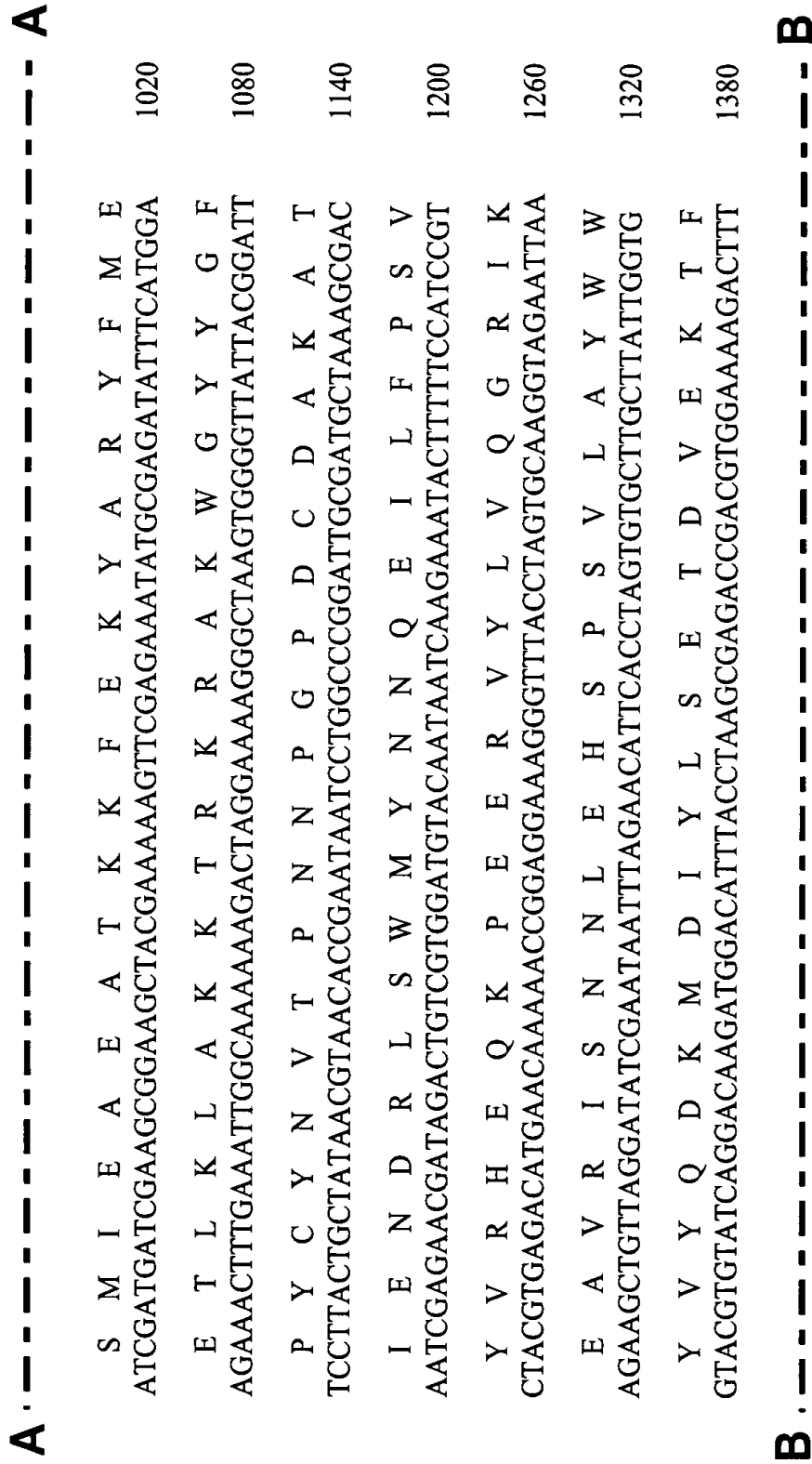
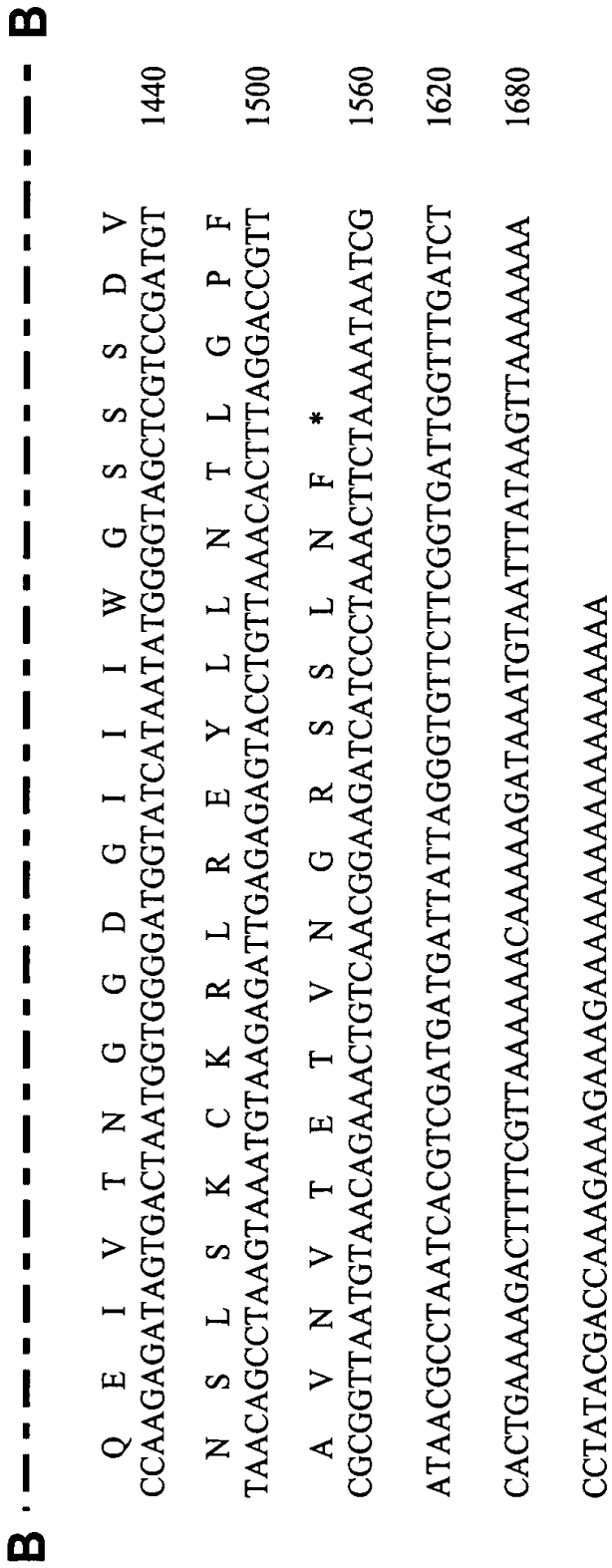


Figure 4B



Note: The amino acid sequence YVSLSP.....RSNCER represents a portion of the leader sequence as the venom protein has the sequence of SKRPKR.....

Figure 4C

pahya, intron sequence, (between nucleotides 733 and 734):

ATTTTCTACTACAGTTCTTTTATCTCTCTATCATTCATTGATGATAAATCGTTTAAATCGAT

60

CTATTGTAAATTATCTATCGATTGTTAGGCAA 94

Figure 5

Vespid hyas:

wfh vv pa bv	1	SERPKRVFNI	YWNVPTFMCH	QYGLYFDEVT	N.FNIKHNSK	50	DDFQGDKISI
		SERPKRVFNI	YWNVPTFMCH	QYDLYFDEVT	N.FNIKRNSK		DDFQGDKIAI
		SERPKRVFNI	YWNVPTFMCH	QYGMNFDEVT	D.FNIKHNSK		DNFRGETISI
		NNKTVREFNV	YWNVPTFMCH	KYGLRFEFVS	EKYGILQNWM		DKFRGEEIAI
wfh vv pa bv	51	FYDPGEFPAL	LPLKEGNYKI	RNGGVPQEGN	ITIHLQRFIE	100	NLDKTYPNRN
		FYDPGEFPAL	LSLKDGYKK	RNGGVPQEGN	ITIHLQKFIE		NLDKIYPNRN
		YYDPGKFPAL	MPLKNGNYEE	RNGGVPQEGN	ITIHLQQFNE		DLDKMTPDKN
		LYDPGMFPAL	LKDPNGNVVA	RNGGVPQLGN	LTKHLQVFRD		HLINQIPDKS
wfh vv pa bv	101	FNGIGVIDFE	RWRPIFRQNW	GNNMIHKKFS	IDLVRNEHPF	150	WDKKMIELEA
		FSGIGVIDFE	RWRPIFRQNW	GNMKIHKNFS	IDLVRNEHPT		WNKKMIELEA
		FGGIGVIDFE	RWKPIFRQNW	GNTIHKKKYS	IELVRKEHPK		WSESMIEAEA
		FPGVGVIDFE	SWRPIFRQNW	ASLQPYKKLS	VEVVRREHPF		WDDQRVEQEA
wfh vv pa bv	151	SKRFEKYARL	FMEETLKLAK	KTRKQADWGY	YGYPYCFNMS	200	PNNLVPDCDA
		SKRFEKYARF	FMEETLKLAK	KTRKQADWGY	YGYPYCFNMS		PNNLVPECDV
		TKKFEKYARY	FMEETLKLAK	KTRKRAKWGY	YGFPYCYNVT		PNNPGPDCDA
		KRRFEKYGQL	FMEETLKAAC	RMRPAANWGY	YAYPYCYNLT		PNQPSAQCEA

A-----A

Figure 6A

A-----A

251					300
wfh	NLKHS.PKVL	SYWWVYVYQDD	TNTFLTETDV	KKTFQEIAIN	GGDGIIIWGS
vv	NLKHS.PKVL	SYWWVYVYQDE	TNTFLTETDV	KKTFQEIVIN	GGDGIIIWGS
pa	NLKHS.PKVL	AYWWVYVYQDK	MDIYLSETDV	EKTFQEIVTN	GGDGIIIWGS
bv	QMTTSRKKVL	PYYWYKYQDR	RTDLSRADL	EATLRKITDL	GGDGIIIWGS
301					
wfh	SSDVNSLSKC	KRLREYLLTV	LGPI TVNVTE	TVN~~~~~	~~~~~
vv	SSDVNSLSKC	KRLQDYLLTV	LGPIANINVTE	AVN~~~~~	~~~~~
pa	SSDVNSLSKC	KRLREYLLNT	LGPF AVNVTE	TVNGRSSLNF	~~~~~
bv	SDDINTKAKC	LQFREYLNNE	LGPA VKRIAL	NNANDRLTV	DVSVDQV*

Figure 6B